

# Discriminação racial, de gênero e conta imigrantes no acesso ao mercado de trabalho

André Spíndola Xandó Baptista, Ricardo Semião e Castro

EESP - FGV

Junho 2020

## Agradecimentos

Gostaríamos de agradecer em especial a todos os professores desta disciplina que tiveram um esforço extra em nos auxiliar no entendimento da matéria, de maneira não usual pelo meio online, e em particular nossa tutora Gabriela Fonseca que se dispôs em diversos momentos a tirar nossas dúvidas e contribuiu com *feedbacks* que nos foram de grande ajuda. Além deles, agradecemos a nossos colegas de sala que colaboraram com debates em aula, principalmente na explicação de conceitos abordados nas bibliografias dos problemas do *workbook* de econometria.

## Resumo

Dada a vasta quantidade de estudos sobre discriminações dentro do mercado de trabalho, gostaríamos de contribuir com a literatura do tema analisando se dita discriminação também ocorre com indivíduos que tentam entrar neste mercado. Dessa forma, pretendemos examinar o efeito da raça, gênero e nacionalidade no tempo em que um indivíduo demora para conseguir um emprego. Para isso, realizamos regressões *cross-section* em uma base de dados dos Estados Unidos incluindo somente pessoas que estiveram desempregadas no ano de 2016. Nos resultados obtidos, encontramos impactos significativamente negativos para mulheres e imigrantes; já para negros, significativamente positivos.

**Palavras chave:** Desemprego, Estados Unidos *cross-section*, código JEL: J

## 1 Introdução

A recente onda de protestos pela igualdade racial nos EUA gerou manifestações no mundo todo, motivando lutas por igualdade em outros aspectos, como movimentos contra o machismo e a xenofobia. Uma das questões mais importantes presentes em qualquer debate sobre desigualdades se refere às causas e pilares de sustentação do fenômeno, porque, se desejamos combater esses contrastes injustos entre certos grupos, precisamos primeiramente entender porque eles existem e como se mantém.

As causas das narrativas preconceituosas citadas - o berço patriarcal do machismo e as raízes coloniais do racismo - são detalhadas pela história e sociologia, bem como já existem inúmeros

estudos que coletaram evidências estatísticas das desigualdades entre homens/mulheres e brancos/negros em vários âmbitos diferentes. A questão mais contundente é então analisar o que perpetua as desigualdades, isso é: buscar quais fatores - que são negativamente afetados pelo preconceito - contribuem para a marginalização de certos grupos, gerando então uma sistematização deste processo.

Para tentar definir os fatores descritos no parágrafo anterior, o presente trabalho utilizará informações sobre o intervalo de tempo que uma pessoa fica desempregada durante um ano como uma métrica para marginalização de um indivíduo. A justificativa do uso dessa estatística se dá não somente pelos problemas financeiros que o desemprego causa, mas também pela diminuição de relações sociais na esfera do trabalho. Além disso, a maior parte da literatura observou evidências das desigualdades pela taxa de desemprego, tentaremos então ver se o mesmo ocorre no âmbito ligeiramente diferente de nossa análise.

Ademais, acreditamos que existam pessoas que sejam extremamente prejudicadas por estarem em dois grupos potencialmente discriminados ao mesmo tempo. Por esse motivo, iremos realizar algumas iterações entre esses grupos a fim de medir o impacto dessas discriminações conjuntas, de modo a mapearmos outros conjuntos de indivíduos que também estejam sofrendo uma grande discriminação e nossas intuições iniciais não os tenham abrangido nesse aspecto.

Sendo assim, nosso objetivo será responder se características como sexo, raça e nacionalidade, afetam causalmente no tempo desempregado, através de um modelo de regressão múltipla. Adicionalmente tentaremos elencar quais são os grupos que mais sofrem com a marginalização, através de interações entre as variáveis.

Revisando a literatura sobre o tema, encontramos evidências fortes que negros possuem grande desvantagem em relação à brancos em processos seletivos de empregos, e os dados da tabela 1 (formulação própria, com dados de nossa base de escolha, a *SIPP*) nos fazem pensar que essa desvantagem pode estar contribuindo para a reentrada mais demorada de negros no mercado de trabalho. Para os imigrantes, os *papers* apontam que os efeitos de discriminação e de falta de conhecimentos sobre o país são os que mais os prejudicam a conseguir um trabalho. Já para as mulheres encontramos indícios tanto para um tempo de desemprego maior em relação aos homens, tanto para um tempo menor; este último parece se sobressair, dado que a tabela 1 aponta que mulheres em média ficam desempregadas duas semanas a menos do que os homens. No entanto sabemos que esses dados somente nos trazem indícios e não nos explicam muita coisa.

Características		Valor	Características		Valor
Sexo	Homem	26	Local de Nascimento	EUA	26
	Mulher	24		América	24
Raça	Branco	25		Africa	25
	Negro	28		Ásia	24
	Asiático	25		Europa	25
	Outra	24		Outros	33

Tabela 1: Média de semanas desempregado por característica

Ao fim do nosso trabalho, obtivemos tanto evidências que suportam os resultados atingidos na literatura (para raça), quanto que suportam o contrário (para sexo e nacionalidade). O resultado da interação entre variáveis evidencia que mulheres negras e homens pobres possam ser grupos mais vulneráveis, bem como ajuda a explicar os resultados contraintuitivos mencionados. Por último, a análise de nacionalidade foi prejudicada e conseguimos apenas resultados “fracos”.

O *paper* está organizado da seguinte maneira: na próxima seção (2) iremos rever a produção literária relevante à nossa pesquisa; na terceira seção iremos apresentar a base de dados utilizada e motivar as variáveis usadas em nossa análise; na quarta seção será explicitada a metodologia utilizada, acompanhada dos resultados de cada modelo criado; em seguida discutiremos esses resultados e as hipóteses levantadas, e por fim traremos uma conclusão do estudo.

## 2 Revisão de bibliografia

Geralmente os artigos já presentes na literatura focam no estudo de algum tipo específico de minoria/discriminação e em nosso trabalho queremos fazer uma pesquisa mais global. Por esse motivo, vamos estudar os três grupos já citados em 3 subseções (uma para cada grupo).

Além disso, não encontramos artigos tratando diretamente sobre o tempo de desemprego, mas sim outras métricas em relação à entrada no mercado de trabalho, as quais julgamos serem bastante representativas da nossa variável em análise.

### 2.1 Gênero

A relação entre desemprego e gênero já foi extensivamente estudada, e a maior parte da literatura tenta evidenciar o *gender unemployment gap* e explicar como ele se comporta com o passar do tempo (SAHIN, 2017) ou em diferentes países (AZMAT, 2004), as explicações dos respectivos artigos não são tão importante para nosso *paper*, mas as evidências coletadas da existência do *gap* nos levam a crer que o gênero provavelmente afeta o tempo desempregado.

Já a análise sociológica presente no artigo “*Unemployed: What Men’s And Women’s Divergent Experiences Tell Us About Gender Inequality*” (2016) de Aliya Hamid Rao nos dá melhores insumos para entender quais são os fatores por trás do gênero que influenciam nossa variável de interesse. Sua principal conclusão é o maior incentivo à busca de emprego que homens recebem da família ao ficarem desempregados, em relação a mulheres: Os homens acabam se sentindo “*trespassers in their own home, who do not belong there during working hours*”, recebendo uma atenção especial que os permite não participar das tarefas domésticas, algo que normalmente não ocorre quando é a mulher que fica desempregada, esta fica dividida entre procurar emprego e as tarefas normalmente atribuídas às “doras de casa”. O *paper* cita que a análise não se aplica apenas à mulheres casadas, também para uma filha em relação à seu pai por exemplo. Entretanto, em famílias onde o homem desenvolve um senso de empatia maior com a mulher o efeito descrito é diminuído.

Por último, analisamos um *paper* de uma visão mais econômica - que gera algumas explicações para porque mulheres poderiam ficar menos tempo desempregadas, em oposição ao artigo de Hamid - o “*Unemployment and wellbeing: unpacking gender differences*” (LONGHI, 2017) tenta explicar através de uma regressão linear como as diferenças de gênero impactam o efeito do desemprego na satisfação de vida, usando dados da “UK Household Longitudinal Survey”. artigo atesta os resultados obtidos em outros estudos e vai de acordo com a intuição de Hamid ao encontrar que na escala de 1 (completamente insatisfeita) a 7 (completamente satisfeita) o coeficiente de perder o emprego condicional à ser homem era  $-0,3$ , enquanto o condicional à ser mulher era aproximadamente  $-0,2$  (após controlado por outras variáveis como tipo de trabalho que fora perdido).

Em segundo lugar o estudo fez sua análise condicional ao nível de importância que o indivíduo dá para pautas igualitárias (contínuo, em porcentagem, onde 0% é ser totalmente tradicional e 100% totalmente igualitário), e obteve que a mulheres igualitaristas sofrem muito mais com o

desemprego que as totalitárias, e como essa discrepância não é tão alta para homens, a partir de um dado valor de igualitarismo as mulheres passam a sofrer mais (resultados na figura 1 no apêndice A), indicando um motivo de porque mulheres poderiam ter maior motivação para buscar empregos, e seus cônjuges maior motivação em apoiá-las, criando uma relação com o último ponto feito sobre o artigo de Hamid.

Em terceiro lugar, foi testado se o efeito descrito anteriormente difere entre mulheres solteiras e casadas, e entre mães e não mães (resultados obtidos estão na figura 2 no apêndice A), motivando que testemos se existe alguma relação entre gênero e estado civil, e gênero e ter filhos. Um último resultado adicional é o de que personalidade não influencia na perda de satisfação advinda do desemprego, que gera maior robustez para a hipótese de não correlação entre os erros e as variáveis do nosso modelo, dado que não teremos essa variável em nossa base de escolha.

## 2.2 Raça

O paper “*Are Emily and Greg More Employable Than Lakisha and Jamal?*” de Marianne Bertrand apresenta um estudo sobre uma possível descriminação racial no acesso ao mercado de trabalho. Mais especificamente, busca se há alguma relação entre a raça de um indivíduo (identificada pela *proxy* de nome que soa “extremamente afro americano” ou “extremamente branco”) e a chance de ele obter um retorno o convocando para uma entrevista presencial após submeter seu currículo para empresas que estavam anunciando vagas de emprego nos jornais.

Os pesquisadores enviaram 4 currículos para cada anúncio, 2 considerados de qualidade alta e 2 considerados de qualidade baixa e aleatoriamente atribuíram um nome soando afro americano para um dos de qualidade alta e um dos de qualidade baixa.

Dentre os resultados, é apresentado que os *White names* recebem em média 50% mais retornos que os *African-American names*. Também encontramos que a raça afeta os benefícios de se ter um bom currículo. Para nomes brancos, um CV (*curriculum vitae*) de alta qualidade provoca 30% mais *callbacks*; para negros, esse valor cai para 9%, no entanto não é estatisticamente significante.

Mais um *outcome* desse estudo é interessante por apontar evidências de descriminação. Uma das variáveis que ajudavam a definir se o currículo era de alta ou baixa qualidade era a *employment holes* (binária), que assumia valor 1 caso a pessoa ficasse um período de pelo menos 6 meses sem reportar um emprego, e 0 caso contrário. Após as regressões serem feitas, essa variável aparecia prevendo positivamente os retornos recebidos, o que os pesquisadores apontaram ser um resultado contra intuitivo. Ainda assim, ela previa mais *callbacks* para brancos do que para negros.

## 2.3 Nacionalidade

No que tange a empregabilidade de imigrantes, estudamos o paper “*Why does unemployment differ for immigrants?*” de Stephen Drinkwater. Ele introduz a ideia de que características pessoais e aspectos socioeconômicos (fatores que afetam a probabilidade de ocorrência de desemprego) frequentemente, mas nem sempre, afetam os nativos e estrangeiros de maneiras semelhantes. Isso implica que há heterogeneidade no impacto quantitativo das características pessoais e de grupo.

O “tipo” de imigrante (legalmente admitidos, refugiados, residentes temporários, como estudantes e trabalhadores temporários e imigrantes sem documentos), bem como seu país de origem, englobam coisas do tipo idioma, cultura e religião, e podem afetar o comportamento desses indivíduos, por exemplo em relação aos tipos de ocupações que podem ser consideradas apropriadas,

bem como influenciar os níveis de discriminação e preconceito.

O artigo levanta também que o capital humano tem uma grande influência na chance de um indivíduo ser empregado. Assim, afirma que as taxas de desemprego variam consideravelmente para os imigrantes, dependendo de suas habilidades e qualificações. Isso ocorre não apenas por causa do vínculo convencional entre educação e probabilidade de emprego, mas também porque o capital humano pode não ser transferível através das fronteiras nacionais. Em particular, os novos migrantes tendem a estar em desvantagem em comparação com os residentes nativos no que diz respeito às habilidades específicas do país (incluindo um conhecimento insuficiente dos costumes do país anfitrião e menor probabilidade de possuir licenças profissionais relevantes) e informações sobre oportunidades no mercado de trabalho. Além de outras habilidades específicas de cada país, a fluência no idioma principal do país anfitrião é um determinante extremamente importante para o sucesso do mercado de trabalho.

Por fim, o artigo aponta que todas essas evidências destacam o papel crucial que o capital humano desempenha na determinação das taxas de desemprego dos imigrantes. No entanto, os imigrantes também podem encontrar discriminação no mercado de trabalho, o que significa que seu capital humano pode não ser recompensado tão favoravelmente quanto os nativos. Em outras palavras, interações entre ser ou não ser imigrante e níveis de educação/fluência do idioma/habilidades preveem um menor efeito de maior capital humano na empregabilidade para imigrantes em relação aos nativos, reduzindo assim suas perspectivas de emprego.

### 3 Dados

#### 3.1 Base de Dados

Para analisar essa relação, escolhemos usar os dados da *SIPP - Survey Income and Program Participation*, uma pesquisa dos Estados Unidos, que de acordo com o site oficial do *Census Bureau* tem seu foco em:

*“The survey’s mission is to provide a nationally representative sample for evaluating: 1) annual and sub-annual income dynamics; 2) movements into and out of government transfer programs; 3) family and social context of individuals and households; and 4) interactions among these items.”*

A base escolhida é interessante pois leva em conta o contexto estadunidense - palco dos protestos referidos na introdução - e contém as variáveis de desemprego (dentro do ponto 1 da citação) bem como as características que queremos (dentro do ponto 3), além de outras variáveis que precisamos incluir para lidar com problemas de endogeneidade.

Iremos utilizar os dados coletados em 2016, onde aproximadamente 230.000 pessoas foram entrevistadas, porém eliminando o número de pessoas que não ficaram desempregadas sobram n=15.907 observações.

#### Amostragem

A secção de amostragem do site oficial da SIPP e o capítulo 7.1 do guia de usuário descrevem como é feita a amostragem da SIPP. A pesquisa faz um ajuste na probabilidade de cada indivíduo ser selecionado para tentar obter um número de indivíduos de cada grupo mais parecido (i.e.: como ocupam uma parcela menor da população, negros tem uma probabilidade maior de serem selecionados), porém a entrevista presa por ter uma amostragem próxima de i.i.d. e faz com que esse ajuste seja mínimo, tentando selecionar com probabilidade ao menos parecida cada indivíduo

residente dos EUA, permitindo-nos manter a hipótese de amostra i.i.d. sem grandes perdas de generalidade. Por outro lado, não teremos observações de estrangeiros suficientes para estimar todos os nossos modelos.

Algumas das possíveis causas de viés e erro de medida seriam: algum agrupamento de pessoas desistirem da entrevista (após a seleção aproximadamente aleatória) com maior probabilidade; algum agrupamento ser escolhido mais frequentemente, uma vez que pessoas podem “aprender” a entrevista e respondê-la “no automático”; algum agrupamento mal interpretar as perguntas com maior frequência que outros grupos. A última causa é mitigada pelo cuidado de escrever perguntas da maneira mais clara possível, e para as duas primeiras não existem evidências de que algum grupo tenha as tendências sistemáticas descritas, também assumimos que a base não incorreu em nenhuma falha técnica que gerasse erro de medida.

### 3.2 Variáveis

#### Variável explicada - semanas desempregado

A base encaixa alguém na categoria desempregado apenas se procurou ativamente emprego por uma ou mais semana por mês. Para quem se encaixa nessa categoria, temos dados da duração de até quatro períodos onde o indivíduo ficou desempregado, além de especificar se o período começou antes de 2016.

Para montar a variável explicada escolhemos: 1. Eliminar períodos que não se iniciaram no ano para especificar o universo de pessoas que analisamos e consequentemente diminuir a possibilidade haver variáveis não observadas que afetassem apenas períodos muito grandes de desemprego (como desmotivação advinda de grandes períodos de procura por exemplo); e 2. Usar o maior período (entre os possíveis 4 apresentados) assumindo que períodos muito pequenos são mais afetados por fatores que não queremos observar (como o hiato entre trocas de emprego por exemplo). No nosso modelo a variável foi chamada de *DESE* como em “semanas em desemprego”.

#### Variáveis explicativas - objetivo

As variáveis principais, que motivam o artigo e sobre quais queremos testar nossas hipóteses discutidas na introdução são:

- Raça autoafirmada - separada entre as *dummies*: branco *RAÇAb*, negro *RAÇAn*, asiático *RAÇAa*, e outras raça *RAÇAo*;
- Sexo - representada pela *dummy* de ser do sexo feminino *SEXOf*;
- Nacionalidade - separada entre as *dummies*: local *NACIlc*, americano (menos nascidos nos EUA) *NACIam*, africano *NACIaf*, asiáticos *NACIas*, e outras nacionalidades *NACIou* - Uma das limitações de nossa base é que a nacionalidade separada por continentes não é tão precisa, não consegue capturar se existe maior preconceito entre canadenses e mexicanos por exemplo.

#### Variáveis explicativas - controle

Nossas variáveis objetivo são definidas pela natureza, o que é um forte motivo para acreditar que não encontraremos nenhum problema de equações simultâneas, portanto o resto das variáveis tem o objetivo de evitar a outra possível causa de viés, o viés de variável omitida. Ainda assim, o resultado da estimação do coeficientes dessas variáveis pode ser utilizado em nossa análise de fatores

que marginalizam os grupos estudados. Para todas as variáveis explicaremos porque elas tem uma relação com nossa variável explicada e também com alguma de nossas explicativas, justificando a necessidade de inclusão da mesma.

- Nível de inglês - separada entre 4 níveis de inglês cada um com sua *dummy* (onde 1 é fluente e 4 é não saber falar): Níveis 1 a 4 *INGL1* a *INGL4* - Essa variável foi adicionada pois imaginamos que melhores falantes de inglês tem menos dificuldade em encontrar emprego, e também que haveria uma relação com ser imigrante ou não.
- Educação concluída - separada em 5 níveis: 4<sup>a</sup> série ou menos, 5<sup>a</sup> à 9<sup>a</sup> séries, colegial, graduação, e mais que graduação, representados respectivamente pelas *dummies EDUC1* a *EDUC5* - A relação entre educação e desemprego já foi amplamente evidenciada na literatura, e além disso, de acordo com (HAMMOND, 1998), o sistema educacional dos EUA é um dos mais desiguais em relação a raça.
- Idade - representada por *IDAD*, variando de 15 anos (idade que já permite um indivíduo exercer trabalhos de rotina não perigosos nos EUA) até 90 (maior idade de um desempregado na base) - Imaginamos que idade tenha relação com as variáveis explicativas por evidências presentes na pirâmide etária norte americana (POPULATION PYRAMID, 2019) por exemplo, que mostra que em média mulheres ocupam em maior proporção as idades mais avançadas. Também incluímos idade ao quadrado para descrever uma possível relação não linear entre idade e tempo desempregado, como já foi apontado em outros estudos estudando taxa de emprego e idade<sup>1</sup>.
- Soma da renda anual, limpa de salário - representado por *RENDl* = renda anual total - salário anual, onde a primeira leva em conta salários/investimentos/seguro social/outros - A renda precisa ser limpa de salário porque este é explicado por semanas desempregadas, logo teríamos viés de equações simultâneas. Incluímos essa variável dada a intuição de que pessoas mais ricas podem se dar ao luxo de ficar mais tempo desempregados, e somado a isso, existem evidências de concentração renda por fatores raciais e de sexo por exemplo<sup>2</sup>.
- Estado de residência - representado por uma *dummy* para cada estado, agrupadas na matriz nx50 *RESI*, e os respectivos coeficientes em outra matriz 50x1 - Pela característica federalista dos EUA, existem muitas diferenças de políticas adotadas entre os estados, as regras do seguro desemprego por exemplo, como apontado na sessão *unemployment* do site oficial do governo federal norte americano. De modo geral, estamos limpando tudo aquilo que um estado pode ter de diferente do outro de maneira sistemática.
- Cidadania - representada pela *dummy* de ser cidadão legal estadunidense *CITI* - Não ter cidadania reconhecida impõe várias restrições legais que impedem um indivíduo de conseguir emprego, e como em nossa base de dados todo nascido nos EUA tem cidadania estadunidense, essa variável é necessária em nosso modelo.
- Uso de recursos voltados para conseguir empregos - a *metadata* da SIPP descreve as seguintes variáveis: 1. “attend any job readiness training to learn about resume writing, job interviewing,

<sup>1</sup>A figura 4 do paper (SIERDZINSKA, 2013) por exemplo retrata o gráfico de o nível de emprego por idade como algo próximo de uma parábola

<sup>2</sup>Em sua análise, (NPWF, 2020) conclui que nos EUA uma mulher negra recebe em média 38% a menos que um homem branco não hispânico

*or building self-esteem?”; 2. “attend any job search program or job clubs, or use any job resource centers to find out about jobs, to schedule interviews, or to fill out applications?” e 3. “attend any training to learn specific job skills, such as computers, car repair, nursing, day care work, or some other job skills?”, que foram unidas na *dummie* (com valor 1 se algum dos programas foi utilizado, 0 caso contrário) *PROG* - É trivial justificar que *PROG* afeta diretamente a capacidade de encontrar emprego, e igualmente sólido assumir que renda influencia na acessibilidade desses recursos, motivando a inclusão dessa variável. Juntamos os três tipos de recursos pois não estamos interessados em identificar qual deles é mais efetivo, apenas evitar viés de variável omitida.*

- Filhos em casa - representado pelas *dummies* (uma para cada intervalo) de ter que sustentar filhos menores de 5 anos *FILH1* e entre 5 e 21 anos *FILH2* - Imaginamos que pais que tem filhos para sustentar não conseguem esperar por um emprego mais desejável e portanto ficariam menos tempo desempregado. Porém, para levar em conta o efeito positivo em *DESE* que “gastar” tempo cuidando de filhos pequenos causa, separamos os dois intervalos descritos para aumentar o poder explicativo do modelo. Além disso, imaginamos que possa haver uma relação cultural e o número de filhos, sendo necessária a inclusão dessa variável.
- Estado civil - representada pela *dummy* de ser casado *ESCVc* - Essa variável tem intuições contrárias: para pessoas casadas que precisam sustentar seu cônjuge devemos observar menos tempo em desemprego, mas quando o cônjuge trabalha, esperamos o efeito contrário. Não esperamos tanta relação com as outras variáveis explicativas, relações do tipo “x raça casa mais” não contam com muitas evidências na literatura, porém, iremos analisar o efeito que ela gera na estimação dos coeficientes para tentar entender se ela seria necessária ou não.
- Tipo da região de residência - representada pela *dummy* de região metropolitana *METR* - Como existem estados mais ou menos rurais, se mercados de trabalho de tipos de regiões diferentes forem mais saturados (ou afetarem o tempo desempregado por qualquer outro motivo) precisaremos incluir essa variável. Também não temos evidências tão fortes para a necessidade dessa variável e iremos analisar o efeito de incluí-la em nossos coeficientes.

## Variáveis explicativas - omitidas

Existem algumas variáveis que imaginamos que podem gerar viés de variável omitida, porém, infelizmente, não estão disponíveis. Iremos discutir qual seria o sinal do viés que elas causariam para entender quais testes estatísticos podem estar comprometidos.

- Religião - Essa variável seria outra como *CITI* que ajudaria a entender melhor como é a influência real de ser imigrante em nossa variável explicada: boa parte da discriminação contra imigrantes árabicos é provavelmente explicada por discriminação contra muçulmanos, por exemplo. A omissão de religião provavelmente não gera problemas no coeficiente de *NACam*, dada a maioria religiosa de matriz cristã e ao contrário, provavelmente gera uma superestimação (em módulo) em *NACas* e possivelmente também em *NACaf*, porém para outros continentes, a variedade de religiões é muito grande e é complicado entender o sinal do viés.
- Tipo de imigrante - Novamente para melhor entender os efeitos da nacionalidade, é preciso controlar pelo motivo do estrangeiro estar nos EUA, como discute o Drinkwater em seu artigo, o “tipo” de imigrante altera sua relação com o desemprego.

- Profissão aplicada - Indivíduos com níveis diferentes de educação (por exemplo) focam em áreas diferentes de atuação, e como a saturação do mercado muda para o tipo de profissão, podemos estar capturando o efeito dessa saturação em algumas de nossas explicativas (*EDUC*, *INGL*, *NACI*). Para entender o sinal do viés associado a essa omissão, precisaríamos entender quais profissões estão associadas a cada um dos níveis de *EDUC* e *INGL* e então a saturação média do mercado média (idealmente ponderada pela popularidade da profissão entre pessoas de um mesmo nível) para essas profissões. Infelizmente essa análise está além do nosso trabalho, indicando uma questão em aberto para as próximas pesquisas no tema.
- Cônjugue trabalha? - Esta variável pode capturar em parte o nível de motivação do indivíduo de procurar um emprego, já que se há outra pessoa arcando com os gastos da família, a urgência de uma renda adicional pode ser menor, e obviamente tem uma relação com *ESCVc*.
- Número de currículos enviados - Essa variável seria uma possível *proxy* para o efeito da motivação de conseguir um novo emprego, uma variável não observada em nosso modelo.

## 4 Metodologia e Resultados

### 4.1 Controles e Modelo Geral

Para entender melhor os efeitos das variáveis sobre *DESE*, e quais controles são necessários, iremos definir nosso modelo inicialmente apenas com as “variáveis explicativas - objetivo” e ir adicionando as “variáveis explicativas - controle” aos poucos, de modo a entender como os coeficientes de raça, sexo e nacionalidade são afetados. As estimações foram feitas com correção de White, dado que nosso modelo mostrou evidências de heterocedasticidade (como será discutido na seção de testes). A comparação dos resultados está na tabela 2 no apêndice B. O primeiro modelo é dado por:

$$DESE = \beta_0 + RAÇA\beta_1 + SEXOf\beta_2 + NACI\beta_3 + \epsilon_1 \quad (1)$$

Onde, para facilitar a notação:

- RAÇA = (RAÇAn | RAÇAa | RAÇAo);
- NACI = (NACIam | NACIaf | NACIas | NACIou)<sup>3</sup>;

Analisando os resultados da tabela 2, vemos que a estimativa retornou um R baixíssimo (0.008), o que sugere que realmente temos variáveis relevantes omitidas, ademais é interessante notar que os efeitos de *SEXOf* (negativo) e *RAÇAn* (positivo) já são significantes .

Começaremos adicionando os controles que temos uma justificativa mais forte para a necessidade, e com os quais imaginamos haver um efeito de maior valor na variável explicada, são eles: estado de residência, educação, idade e nível de inglês, de modo que o modelo 2 fica:

$$DESE = \beta_0 + RAÇA\beta_1 + SEXOf\beta_2 + NACI\beta_3 + RESI\beta_4 + \epsilon_2 \quad (2)$$

Comparando os resultados, percebemos que *RESI* é responsável por explicar boa parte dos efeitos: o R ajustado é aproximadamente 17 vezes maior que o anterior, *SEXOf* e *RAÇAn* estavam provavelmente superestimados (em módulo) e *NACAf* ganhou significância.

Como imaginamos que as variáveis de nível de inglês e educação afetam de maneira parecida

---

<sup>3</sup>Como omitimos as *dummies* de raça branca e nacionalidade local, podemos interpretar os coeficientes como uma diferença de média entre os grupos

nosso modelo (influenciam o nível de capital humano do indivíduo), seguimos adicionando ambas em (3):

$$DESE = \beta_0 + RAÇA\beta_1 + SEXOf\beta_2 + NACI\beta_3 + RESI\beta_4 + INGL\beta_5 + EDUC\beta_6 + \epsilon_3 \quad (3)$$

Onde:

- INGL = (INGL2 | INGL3 | INGL4);
- EDUC = (EDUC2 | EDUC3 | EDUC4 | EDUC5).

Observando os resultados, o coeficiente para homens e para negros novamente diminuíram, mas consideravelmente menos do que do modelo (1) para (2), e notamos que as *dummies* de educação parecem ter maior impacto em relação às de nível de inglês (maior significância e valor absoluto). *NACIeu* ganhou significância no modelo. Em (4), adicionamos a idade e idade ao quadrado:

$$\begin{aligned} DESE = \beta_0 + RAÇA\beta_1 + SEXOf\beta_2 + NACI\beta_3 + RESI\beta_4 + INGL\beta_5 + EDUC\beta_6 \\ + IDAD\beta_7 + IDAD^2\beta_{72} + \epsilon_4 \end{aligned} \quad (4)$$

Idade também influencia bastante o modelo, o intercepto e os coeficientes de *RESI* foram profundamente alterados. Ademais, *RAÇAa*, *NACIam* e *NACIas* se tornaram variáveis significativas, com sinais positivo, negativo e negativo respectivamente. Percebemos que mais uma vez os estimadores de *RAÇAn* e *SEXOf* tiveram um decréscimo em módulo (0,34 e 0,136 nessa devida ordem) o que vale também para os estimadores de *NACIaf* e *NACIeu*.<sup>4</sup>. Agora adicionaremos todos os outros controles que não tínhamos tantas evidências da necessidade, gerando o modelo (5):

$$\begin{aligned} DESE = \beta_0 + RAÇA\beta_1 + SEXOf\beta_2 + NACI\beta_3 + RESI\beta_4 + INGL\beta_5 \\ + EDUC\beta_6 + IDAD\beta_7 + IDAD^2\beta_{72} + EXTRAS\beta_8 + \epsilon_5 \end{aligned} \quad (5)$$

Onde:

- EXTRAS = (RENDI | ESCVc | FILHO1 | FILHO2 | CITI | METRr | METRu | PROG).

A diferença entre as estimativas de (5) e (4) são menores no geral do que nas outras passagens, para checar se a adição de *EXTRAS* é importante, fizemos um teste F com  $H_0$ : todos os elementos de  $\beta_8 = 0$ , obtendo um p-valor < 0,05, logo as variáveis ajudam a obter um modelo menos endógeno e por isso este será nosso modelo geral definitivo.

Nesse modelo final, encontramos que *RAÇAn* possui efeito positivos no número de semanas desempregado, já para *SEXOf*, *NACIam*, *NACIaf* e *NACIas* os efeitos são negativos (todos esses resultados são estatisticamente significantes). Dois resultados secundários: renda, por mais que significativa, tem um coeficiente baixíssimo de 0,0008; Casados ficam em média 2,45 semanas a menos desempregados (também significante).

## 4.2 Mapeando piores casos

Definidos as nossas variáveis controles, para facilitar a notação iremos agrupar em uma matriz X, onde cada coluna é uma variável de controle. Nossa análise agora consiste em permitir interações entre as variáveis na tentativa de encontrar quais são os grupos da sociedade que mais

---

<sup>4</sup>Após incluir idade, os coeficientes da variável de residência se aproximaram em valores absolutos, indicando que parte dos efeitos de *RESI* eram explicados pela diferença de idade média dos estados.

são afetados por discriminações, para isso o melhor seria relacionar as três variáveis objetivo como no modelo (6). Para facilitar a notação, agrupamos todas as combinações entre *RAÇA*, *NACI*, *SEXO* (lembrando que o grupo omitido de cada se mantém omitido nas interações), multiplicadas pelos respectivos coeficientes, no termo *INT*.

$$DESE = \beta_0 + RAÇA\beta_1 + SEXOf\beta_2 + NACI\beta_3 + INT + X\beta_X + \epsilon_6 \quad (6)$$

Infelizmente, como foi citado na seção de Base de dados, nossa base não tem tantas observações de estrangeiros, tornando impossível estimar (6) - não temos um número de indivíduos suficientes que nasceram na África e se indicaram como asiáticos (raça) para estimar o coeficiente associado à essa interação por exemplo, o problema é interagir raça e nacionalidade. As interações com sexo e ambas as outras felizmente ainda são estimáveis, então começaremos com essa análise no modelo 7. Por facilidade de notação representamos todas as combinações entre as duas variáveis - multiplicadas pelos seus respectivos coeficientes - pelo termo *INT<sub>RAÇA</sub><sup>SEXO</sup>*, essa notação será usada no resto deste artigo.

$$DESE = \beta_0 + RAÇA\beta_1 + SEXOf\beta_2 + NACI\beta_3 + INT_{RAÇA}^{SEXOf} + X\beta_X + \epsilon_7 \quad (7)$$

A análise dos resultados das interações serão apresentadas em formato gráfico para facilitar a visualização da diferença entre os grupos, e somente se houver alguma diferença interessante nos outros coeficientes (em relação aos resultados de (5)) estes serão apresentados.

A figura 4 do apêndice C mostra a diferença entre ser homem para cada raça diferente, conseguimos observar dois resultados principais: apenas as mulheres negras ficam mais tempo desempregadas que homens da mesma raça, e discrepância entre homens e mulheres é maior para brancos. Os coeficientes de asiáticos e de outros não eram significantes.

Dada a motivação presente na introdução de encontrar os fatores que podem perpetuar a marginalização de certos grupos, iremos observar qual é o efeito da renda e da idade intra-grupos, começamos pelas interações com IDAD e IDAD (ambos representados por IDAD na notação proposta anteriormente):

$$DESE = \beta_0 + SEXOf\beta_2 + INT_{SEXOf}^{IDAD} + RAÇA\beta_1 + NACI\beta_3 + X\beta_X + \epsilon_8 \quad (8)$$

As funções na figura 5 do apêndice C mostram a variação em *DESE* que a diferença entre os sexos causa, e como essas diferenças são alteradas pela idade<sup>5</sup>. Podemos perceber que aproximadamente a partir dos 63 anos as mulheres começam a ficar mais tempo desempregadas que homens da mesma idade, atingindo uma diferença de 7 semanas para idades mais altas.

Realizando a mesma análise para raça (apenas trocando *INT<sub>IDAD</sub><sup>SEXOf</sup>* por *INT<sub>IDAD</sub><sup>RAÇA</sup>* em (7)), obtemos o relatório da figura 6, onde podemos perceber que a interação não gera muitas diferenças entre brancos e negros (parábolas parecidas, somente deslocadas dado o coeficiente positivo associado à *RAÇAn*), indicando que a interação com idade não deve ser um fator que sistematiza a marginalização desse grupo. Os coeficientes obtidos para *RAÇAo* não foram significantes. Faremos a mesma análise com renda:

$$DESE = \beta_0 + SEXOf\beta_2 + INT_{SEXOf}^{RENDI} + RAÇA\beta_1 + NACI\beta_3 + X\beta_X + \epsilon_8 \quad (9)$$

O indício de que renda seria um fator como o descrito na introdução seria pobres de um certo grupo ficarem mais tempo desempregado em relação aos outros, isso dificultaria a geração de renda

---

<sup>5</sup> *DESE* = *IDAD* $\beta_7$  + *IDAD* $^2\beta_7$  + *SEXOf* $\beta_2$  + *INT<sub>SEXOf</sub><sup>IDAD</sup>* para mulheres, e apenas os três primeiros termos para homens

e então sair da condição de pobreza. Na figura 7 retratamos evidências de que esse fenômeno ocorre com os homens, e não com as mulheres; Na figura 8, obtivemos uma análise pouco precisa, com apenas os coeficientes associados a função de brancos estatisticamente significativa, mesmo assim, o gráfico indica que para ricos negros e brancos ficam pior que as outras raças, pobres de “outras” ou “asiáticos” tem maior dificuldade de encontrar emprego.

Por último, iremos testar as conclusões do artigo de Longhi, que encontrou evidências sobre o efeito na satisfação após perder o emprego - condicional ao nível de igualitarismo - difere entre mulheres casadas e solteiras, e com filhos e sem. Queremos ver se um fenômeno parecido ocorre em relação ao tempo desempregado, o efeito de ter filhos e ser casado é pior para mulheres do que para homens? Para responder essa pergunta possibilitamos haver interação entre as referidas variáveis em (10)<sup>6</sup>:

$$DESE = \beta_0 + RAÇA\beta_1 + SEXOf\beta_2 + NACI\beta_3 + INT_{SEXOf}^{FILHO} + INT_{SEXOf}^{ESCVc} + X\beta_X + \epsilon_1 1 \quad (10)$$

Os coeficientes estimado para a interação entre ser mulher e ter filhos foi de  $-5,261^{***}$  (0,673); para ser mulher e casada foi 0,632 (0,650). Obtemos então evidências parecidas, para mulheres o efeito de ter filhos é muito maior que o do estado civil, e a diferença no efeito de ter filhos para mulheres e homens também (primeiro resultado) é grande.

Não apresentamos os resultados para interações com *NACI* dada a natureza muito generalista dessa variável, o que gerou coeficientes muito pequenos e pouco significativos na estimação de modelos análogos à (6), (8) e (9).

Para coletar evidências que as interações são significativas em seus respectivos modelos, fizemos um teste F com hipótese nula “os coeficientes de todas as interações presentes no modelo são conjuntamente = 0”, todos os testes tiveram um p-valor  $< 0,05$ , inclusive nos modelos não apresentados de nacionalidade, implicando que uma análise das interações com essa variável ainda é interessante, principalmente com uma outra variável de nacionalidade mais específica.

### 4.3 Testes

Além dos testes F para inclusão de grupos de variáveis em nosso modelos, também fizemos - sobre o modelo final (5) - o teste de White para testar heterocedasticidade e *geometric goodness of fit* entre a estimação de densidade de Kernel para os resíduos e uma normal.

O teste de White gera uma estatística de teste = 631,66, gerando um p-valor  $< 2,2e^{-16}$ , logo rejeitamos a hipótese nula de homoscedasticidade e assumimos heterocedasticidade. Para poder corrigir esse fator - dado que não sabemos o formato da variância e logo não conseguiremos aplicar *Feasible Generalized Least Squares* - iremos estimar todos os modelos com erros robustos de White, para que nossos testes de hipótese possam ser interpretados corretamente.

O histograma dos resíduos  $\hat{\epsilon}_6$ , a estimação de sua distribuição por Kernel, e uma normal com parâmetros iguais à média e desvio padrão da estimativa de Kernel são representados na figura 3 do apêndice 6. O teste *geometric goodness of fit* gera uma estatística de teste = 1,8359 e um p-valor  $< 2.2e^{-16}$ , logo rejeitamos a hipótese nula de distribuição normal.

---

<sup>6</sup>Alteramos a variável *FILHO* para significar apenas “ter filhos em casa”.

O resultado nos leva a crer os testes acerca da significância de nossas variáveis estarão prejudicados.

#### 4.4 Hipóteses de Gauss Markov - Consistência

Como nosso número de observações é grande o suficiente para utilizar os teoremas da análise assintótica, e teríamos problemas em garantir não-viés - dado que temos algumas variáveis omitidas - iremos apenas nos preocupar em garantir a consistência de nossas estimativas, por isso substituímos a hipótese de *exogeneidade* por *ortogonalidade* e de *regressores fixos* por *estabilidade*. Explicitamente, assumimos sobre o modelo (5) que:

- O modelo populacional é linear nos parâmetros;
- Como discutido na seção de Amostragem, iremos assumir que a amostra é independente e identicamente distribuída;
- Não existe colinearidade perfeita entre as variáveis, por mais que existam variáveis com alta correlação (como *RAÇAn* e *NACIaf*) - o que pode aumentar os fatores de inflação da variância - não existe nenhum par de variáveis que tenha colinearidade perfeita, possibilitando a estimação do modelo;
- Estabilidade: assumimos que a matriz de covariância das variáveis explicativas converge em probabilidade para uma matriz definida Q, isto é: ;

Hipóteses que são prejudicadas em nosso modelo:

- Como o número de observações é suficientemente grande, podemos substituir a hipótese de exogeneidade pela mais fraca de ortogonalidade entre as variáveis explicativas  $\epsilon_6$ , pela análise feita na seção de Variáveis, fizemos um extenso trabalho em encontrar controles na tentativa de evitar viés de variável omitida, porém existem algumas variáveis que não tivemos acesso e possivelmente impedem que essa hipótese seja totalmente verdadeira. Nossa intuição sobre as variáveis citadas em “Variáveis explicativas - omitidas” nos leva a acreditar que seus efeitos não são tão significativos em nossa estimação, e que conseguimos encontrar os controles mais importantes, e portanto, no escopo deste trabalho, assumiremos que  $plim(\frac{1}{n}X'\epsilon_6) = 0$  para poder fazer algum tipo de análise causal entre as variáveis;
- Pelo mesmo motivo da exogeneidade, também temos problemas em assumir ausência de autocorrelação entre os erros, uma vez que podemos ter algum fator dentro de  $\epsilon_6$  que gere correlação entre eles, não teremos mais então estimadores *BLUE* pois perdemos em eficiência, fator que também é alterado pela heterocedasticidade discutida a seguir;
- O termo de erro  $\epsilon_6$  é heterocedástico - como discutido na seção de testes - porém a estimação do modelo por *OLS* robusto permite que nossos testes de hipótese ainda sejam válidos;
- A hipótese sobre o termo de erro  $\epsilon_6$  se distribuir como uma normal - como discutido na seção de teste - será prejudicada, implicando que nossos testes de significância tem sua validade enfraquecida.

## 5 Discussão

Negros demoram aproximadamente 2,2 semanas a mais que brancos para conseguir um emprego. Esse resultado vai de acordo com nossa suposição inicial e com a literatura; esperava-se que, como brancos recebem mais retornos das empresas ao enviarem seus currículos, eles conseguissem sair do desemprego antes comparados aos de pele escura.

Em relação aos imigrantes, todas as nacionalidades que possuem estimadores estatisticamente significantes estão afetando negativamente nossa variável em análise. Isto vai contra a intuição apontada na bibliografia de que imigrantes tendem a ter desvantagem para encontrar empregos por terem menos habilidades específicas do país e menos informações sobre o mercado de trabalho. Acreditamos que esses resultados opostos se dão por uma das limitações de nosso trabalho: não estamos conseguindo capturar informações das exigências dos indivíduos com sua nova posição de trabalho, como salários, grau de satisfação e qualidade do emprego. Dessa forma, justamente pelos imigrantes terem ciência da maior dificuldade que terão de encontrar empregos, imaginamos que eles estejam aceitando trabalhar em lugares que os forneçam menores benefícios, agilizando assim o processo para sair do desemprego. Para trabalhos futuros, seria interessante utilizar uma variável que capturasse esses efeitos, bem como uma de nacionalidades separadas por países para conseguirmos observar se há alguma evidência de discriminação com populações de países específicos.

No entanto, não descartamos as características da imigração encontradas na literatura. Podemos notar que o estimador de *NACIaf* aumenta consideravelmente (+1,746) quando adicionamos as variáveis de *EDUC* e *INGL* no modelo. Destarte, grande parte do tempo a menos que um africano ficava desempregado nos EUA pode estar sendo explicado pelo seu nível de escolaridade e a fluência no idioma, o que vai de acordo com o discutido no *paper* de Drinkwater. Sem embargo, nenhuma outra variável objetivo que era significativa (*RAÇAn* e *SEXOf*) teve seus estimadores muito afetados, o que indica que o menor tempo de desemprego para mulheres (em comparação a homens) e maior para negros (em comparação a brancos) não estaria sendo explicado pela escolaridade e fluência do inglês.

Em relação ao sexo, encontramos que mulheres ficam desempregadas aproximadamente uma semana a menos em comparação aos homens. Para justificar esse efeito contra-intuitivo, além do ponto levantado no *paper* de Longhi, nas regressões que permitimos interações entre variáveis obtivemos coeficientes menores (em módulo) para *SEXOf*: para a interação com idade, que gera menos semanas desempregadas para mulheres na maior parte da vida (aproximadamente até os 63 anos), o coeficiente associado à ser mulher caiu para aproximadamente 0,9; além disso o efeito descrito da idade é maior do que na interação com renda, onde observamos o contrário (efeito de ganho de semanas menor em homens para a maioria das faixas de renda). Além disso, o efeito de ter filhos gerou menor tempo desempregado para mulheres do que para homens, o que vai de acordo com as descobertas no *paper* de Longhi.

As evidências de interações não podem ser comparadas diretamente com o coeficiente do modelo (5), porém os efeitos descritos nos ajudam a entender o porque obtivemos esse resultado contra intuitivo, mostrando que um coeficiente negativo associado a ser mulher não implica em discriminações à homens na entrada do mercado de trabalho.

Dentro do objetivo de mapear os piores casos, nos deparamos com a limitação de não conseguir analisar as três variáveis objetivo juntas, impedindo a análise de “qual grupo é pior” motivada na introdução. Obter uma amostra mais representativa de imigrantes e então poder realizar essa

análise seria uma adição positiva à este trabalho. Nossa análise porém conseguiu gerar resultados para análise de interações menores: obtivemos evidências de que apenas mulheres negras ficam mais tempo desempregadas que homens da mesma raça, corroborando com argumentos acerca da profunda discriminação que esse grupo “duplamente discriminado” sofre. Além disso, dentre as pessoas de idades mais elevadas, as mulheres também ficam mais tempo desempregadas que os homens, indicando que sair do mercado de trabalho em idades avançadas pode ser algo definitivo com maior probabilidade parar mulheres do que para homens e que, portanto, a estabilidade financeira de mulheres mais velhas deve ser uma pauta em debates de igualdade. Ao contrário da interação com sexo, não encontramos tanta evidência para a discrepância do efeito de idade em negros e brancos.

Também obtivemos o resultado inesperado que homens pobres sofrem maiores dificuldades que mulheres pobres, evidenciando talvez a existência de um grupo que possa ter sido deixado de lado dentro do debate da discussão de igualdade: será que em famílias pobres as mulheres recebem maior prioridade na busca de emprego que homens, ao contrário do que Hamid descreveu em seu artigo?

A interação de renda e raça se mostrou pouco significativa, sendo difícil concluir um efeito específico como foi possível nas análises com sexo, porém levantamos evidências mais gerais de que existe alguma interação. Somado à isso, nossa problemática de pouco poder interpretativo em interações com *NACI* não nos permitiu checar intuições descritas por Drinkwater acerca do efeito diferente da educação entre imigrantes e locais por exemplo, deixando o objetivo de procurar efeitos que perpetuam a discriminação de grupos específicos em aberto, possivelmente através de pesquisas com variáveis mais específicas (e com uma amostra mais representativa dos diferentes grupos).

Fora o que já foi citado, outra maneira de refinar o presente trabalho seria incluir uma análise de *time series*, para ver como os coeficientes do modelo (5) se comportam com o passar do tempo, para tentar gerar alguma interpretação de se a discriminação vêm diminuindo na atualidade.

## 6 Conclusão

É necessário ressaltar que nossa análise tem suas limitações, o viés de variável omitida impede que tiremos conclusões extremamente definitivas, porém para raça e sexo principalmente conseguimos fazer uma análise bastante completa ao permitir as interações. Para raça, encontramos evidências na mesma direção da discriminação observada pela taxa de desemprego discutida na literatura. Para sexo, conseguimos elencar como a interação dos controles com sexo pode explicar o resultado que não vai na linha mais geral da literatura (porém que não é totalmente descartado por esta).

A análise dos efeitos da nacionalidade foi a parte mais fraca do modelo, ainda assim contribuímos com evidências de que a relação de imigrantes com educação e nível de inglês possivelmente se dá em dois âmbitos: intuímos que por um lado essas duas características devem ajudar a explicar diretamente o tempo desempregado (e uma interação entre elas e nacionalidade pode ser feita em artigos futuros), mas por outro imigrantes podem se organizar em áreas de atuação diferente por causa delas, afetando indiretamente a variável de interesse.

Nosso objetivo de encontrar grupos muito afetados conseguiu evidências fracas, mas relevantes de discriminações “extras” com mulheres negras e homens muito pobres, abrindo portas para novas discussões sobre esses grupos.

## 7 Referências

A. RUZIK-SIERDZINSKA, M. LIS, M. POTOCZNA, M. BELLONI, C. VILLOSIO. NEU-JOBS Working Paper No. 17.2 **Age and productivity. Human capital accumulation and depreciation.** 2013

Bertrand, Marianne, and Sendhil Mullainathan. 2004. “**Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination.**” American Economic Review, 94

Bureau of Labor Statistics, U.S. Department of Labor, Occupational Outlook Handbook, **Foreign-Born Workers: Labor Force Characteristics - 2019**

Brookings. Articles. Unequal Opportunity: Race and Education. Disponível em: <https://www.brookings.edu/articles/unequal-opportunity-race-and-education/> Último acesso em 24/06/2020

DRINKWATER, Stephan. University of Roehampton, UK, and IZA, Germany. **Why does unemployment differ for immigrants?**. 2017

HEIJI, C; DE BOER, P; HANS FRANSE, P; KLOEK, T; VAN DIJK, H. Econometric Methods With Applications in Business and Economics. Oxford University Press, 2004. 787 p.

National Partnership for Women Families, **Quantifying America's Gender Wage Gap by Race/Ethnicity**

Population Pyramid. Pirâmides Populacionais do Mundo desde 1950 até 2100. Estados Unidos. Disponível em: <https://www.populationpyramid.net/pt/estados-unidos/2019/> Último acesso em 24/06/2020

USA GOV. Unemployment.Unequal Opportunity: Race and Education. Disponível em: <https://www.usa.gov/unemployment> Último acesso em 21/06/2020

United States Census Bureau. **2017 SIPP Production SIPP Public Use Metadata Report.** 2019

United States Census Bureau. Programs Surveys. SIPP Introduction History. Disponível em: <https://www.census.gov/programs-surveys/sipp/about/sipp-introduction-history.html> Último acesso em 25/06/2020

United States Census Bureau. Programs Surveys. Sampling. Disponível em: <https://www.census.gov/programs-surveys/sipp/methodology/sampling.html> Último acesso em 25/06/2020

United States Census Bureaus. **Survey of Income and Program Participation 2014 Panel Users' Guide.** 2014

# APÊNDICE A

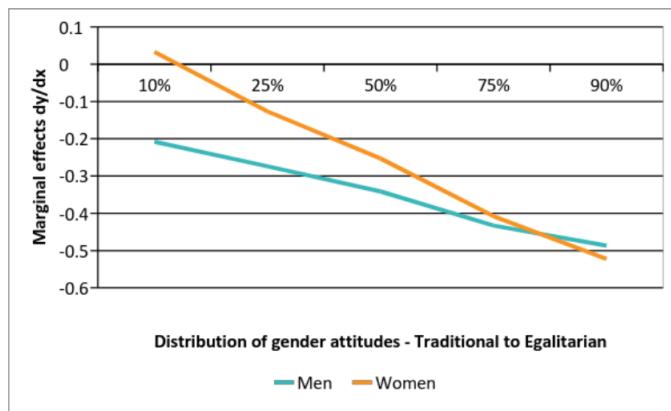


Figura 1: Efeito marginal na satisfação de perder emprego, para homens e mulheres, condicional ao nível de igualitarismo.

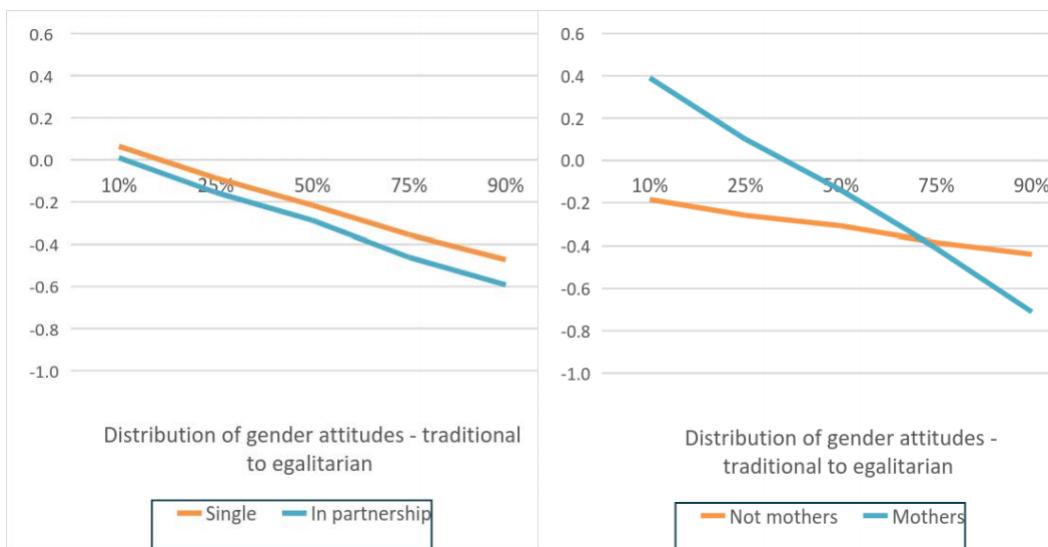


Figura 2: Efeito marginal na satisfação de perder emprego para mulheres, condicional ao nível de igualitarismo, estado civil e ser mãe.

# APÊNDICE B

Tabela 2: Resultado da estimação dos modelos (1) à (5)

	DESE				
	Modelo (1)	Modelo (2)	Modelo (3)	Modelo (4)	Modelo (5)
Constante	23.541*** (0.233)	18.399*** (2.582)	24.996*** (3.579)	6.668* (3.576)	6.436* (2.974)
RAÇAn	3.539*** (0.393)	2.939*** (0.392)	2.811*** (0.390)	2.471*** (0.384)	2.180*** (0.373)
RAÇAa	0.481 (1.063)	-0.668 (1.062)	-0.443 (1.045)	2.190** (1.020)	2.550** (0.872)
RAÇAo	-0.337 (0.578)	-0.230 (0.542)	-0.582 (0.541)	-0.003 (0.532)	0.069 (0.628)
SEXOf	-2.017*** (0.292)	-1.375*** (0.280)	-1.312*** (0.279)	-1.176*** (0.271)	-1.097*** (0.287)
NAClam	0.061 (0.573)	0.060 (0.554)	0.423 (0.608)	-1.148** (0.585)	-0.832 (0.823)
NAClaf	-1.488 (1.559)	-7.005*** (1.543)	-5.259*** (1.559)	-6.511*** (1.523)	-5.088*** (1.395)
NACIeu	0.563 (1.511)	1.906 (1.429)	3.672*** (1.421)	2.832** (1.390)	3.526* (1.575)
NACIas	-0.369 (1.173)	-0.931 (1.180)	-0.692 (1.195)	-3.470*** (1.149)	-2.842** (1.098)
NACIou	-0.915 (0.407)	-0.831 (0.406)	-0.669 (0.396)	1.857* (0.496)	1.847* (0.495)
EDUC2			-7.391*** (2.473)	-2.028 (2.474)	-2.292 (1.921)
EDUC3			-3.995 (2.464)	-1.101 (2.457)	-1.431 (1.903)
EDUC4			-6.534*** (2.469)	-4.825* (2.462)	-5.175** (1.909)
EDUC5			-7.310*** (2.515)	-7.581*** (2.501)	-7.991*** (1.964)
INGL2			-1.111 (0.843)	-1.804** (0.829)	-1.241 (0.924)
INGL3			-2.345** (1.174)	-3.350*** (1.149)	-3.503** (1.341)
INGL4			0.021 (1.759)	-1.314 (1.722)	-1.107 (1.718)
IDAD			0.596*** (0.045)	0.661*** (0.054)	0.661*** (0.054)
IDAD <sup>2</sup>			-0.005*** (0.001)	-0.005*** (0.066)	-0.005*** (0.066)
RESI	Não	Sim	Sim	Sim	Sim
EXTRAS	Não	Não	Não	Não	Não
Observações	15,907	15,907	15,907	15,907	15,907
R <sup>2</sup>	0.008	0.145	0.155	0.214	0.215
R <sup>2</sup> ajustado	0.008	0.141	0.151	0.211	0.201
Erro padrão	0.725	0.723	0.721	0.718	0.716
Resíduo	(df = 15898)	(df = 15847)	(df = 15840)	(df = 15838)	(df = 15830)
Estatística F	16.554***	45.413***	43.964***	63.459***	53.616***
	(df = 8; 15898)	(df = 59; 15847)	(df = 66; 15840)	(df = 68; 15838)	(df = 76; 15830)

Note: \* p<0.1; \*\* p<0.05; \*\*\* p<0.01; RESI e EXTRAS omitidos, indicamos apenas sua presença

## APÊNDICE C

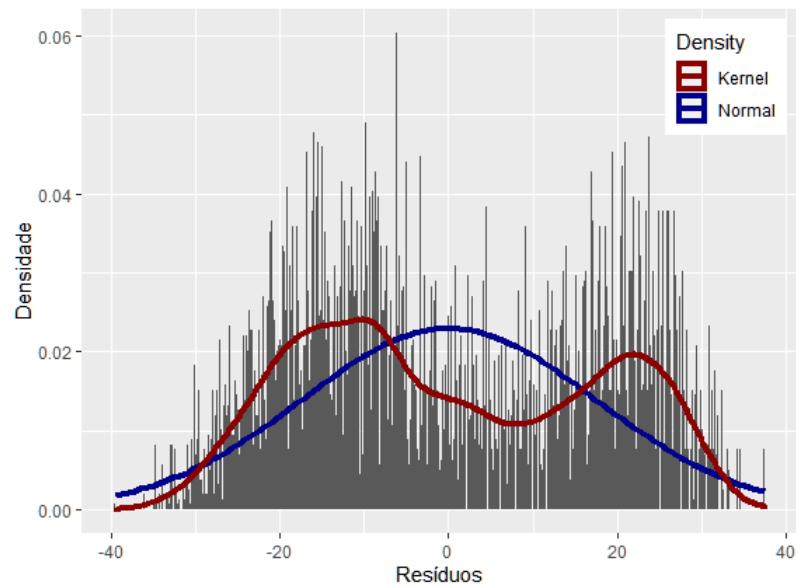


Figura 3: Distribuição de  $\hat{\epsilon}_6$

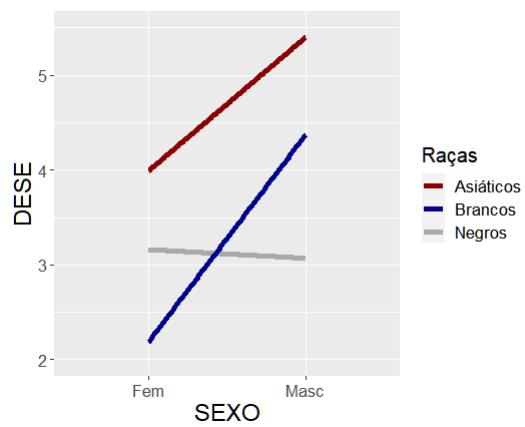


Figura 4: Efeito de SEXO em DESE, por SEXO

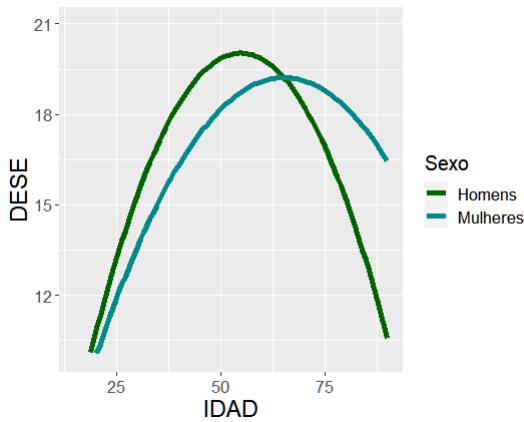


Figura 5: Efeito de IDAD em DESE, por SEXO

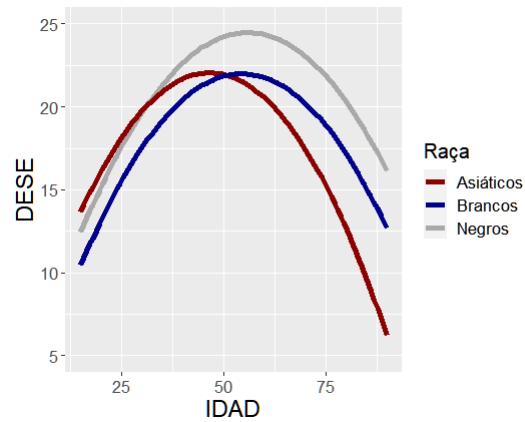


Figura 6: Efeito de IDAD em DESE, por RAÇA

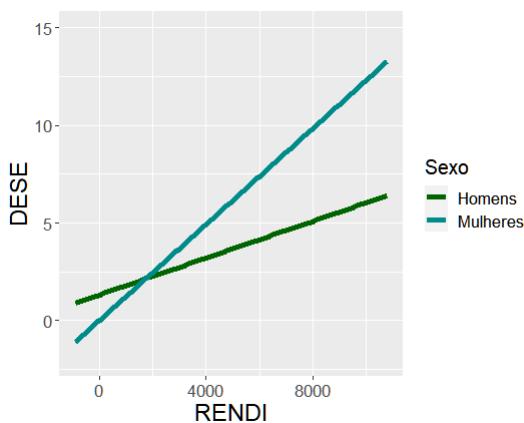


Figura 7: Efeito de RENDI em DESE, por SEXO

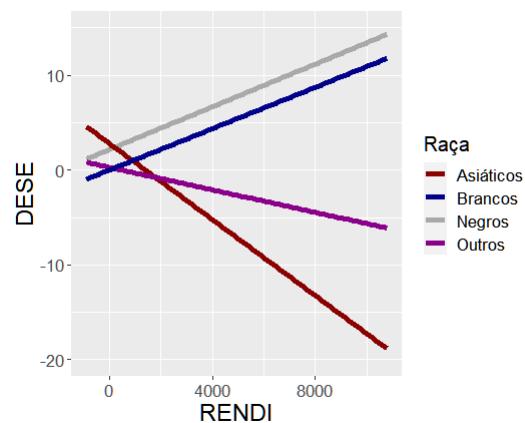


Figura 8: Efeito de RENDI em DESE, por RAÇA